LSR Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 22, 2018

J. Dong
S. Bryant
Huawei Technologies

June 20, 2018

IGP Extensions for Segment Routing based Enhanced VPN
draft-dong-lsr-sr-enhanced-vpn-00

Abstract

Enhanced VPN (VPN+) is an enhancement to VPN services to support the needs of new applications, particularly including the applications that are associated with 5G services. These applications require better isolation and have more stringent performance requirements than that can be provided with traditional overlay VPNs. An enhanced VPN may form the underpin of 5G transport network slicing, and will also be of use in its own right. This document describes how Multi-Topology Routing (MTR) as described in RFC 5120 and RFC 4915, can be extended to signal the resources allocated in the underlay network to construct the virtual networks for enhanced VPN services, together with the Segment Routing Identifiers (SIDs) used to identify and access the network resources allocated for the virtual networks in the data plane.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on December 22, 2018.

Table of Contents

1. Introduction

The framework for an enhanced virtual private network (VPN+) is described in [I-D.bryant-rtgwg-enhanced-vpn].

Driven largely by needs arising from the 5G mobile network design, the concept of network slicing has gained traction. There is a need to create a VPN service with enhanced isolation and performance characteristics. Specifically, there is a need for a transport network to support a set of virtual networks, each of which provides the client with some dedicated (private) network resources drawn from

a shared pool. The tenant of such a virtual network can require a degree of isolation and performance that previously could only be satisfied by dedicated networks. Additionally the tenant may ask for some level of control of their virtual networks e.g. to customize the service paths in their network slices.

These properties cannot be met with pure overlay networks, as they require tighter coordination and integration between the underlay and the overlay network. [I-D.bryant-rtgwg-enhanced-vpn] provides the framework of enhanced VPN and describes the candidate component technologies. [I-D.dong-spring-sr-for-enhanced-vpn] describes how segment routing (SR) [I-D.ietf-spring-segment-routing] is used to construct the required virtual networks with the network resources allocated for enhanced VPN services.

This document describes how Multi-Topology Routing (MTR), as described in [RFC5120] [RFC4915] , is extended to signal the resources allocated in the underlay to construct the virtual networks for enhanced VPN services, together with the segment routing identifiers used to identify and access the resource allocated for different virtual networks in the data plane.

2. Specification of Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

3. Overview of Approach

To meet the requirement of enhance VPN services, a number of virtual networks can be created, each representing a subset of the underlay network topology and resources to be used by a specific customer. In a 5G context, each virtual network is considered as a network slice which serves one slice tenant. Depending on the service requirements, differernt virtual networks can either share the same physical links or nodes, or use separate links or nodes in the network, while the required level of isolation and performance SHOULD be guaranteed in both cases.

IGP multi-topology routing can be seen as a candidate mechanism to create multiple network topologies in one network. Different from the traditional multi-topology mechanism which only provides logical topological isolation, in the proposed mechanism network resources can be partitioned and allocated to different virtual network topologies to meet the isolation and performance requirements of enhanced VPN. Service in one virtual topology can be instructed to be processed using the network resources allocated to this virtual

topology. This is achieved by using multi-topology together with segment routing, and extending the SR paradigm to use Segment Identfiers (SIDs) to identify different set of resources allocated from a particular network element (e.g. link or node). Different set of SIDs are associated with different virtual topologies, and are used to create the SID lists in different topologies. In some cases it is also possible for several virtual network topologies to share some network resources, this can be achieved by using the same SR SIDs between those topologies. The detailed mechanism of resource sharing will be described in a future version.

Within one SR virtual network, one or more type of services can be deployed using the resources allocated to that topology, some of which may have different characteristics and require dedicated resources or special treatment. This is similar to the DS-TE model of the RSVP-TE based mechanism. The concept is similar to the DS-TE model [RFC4124] of RSVP-TE based mechanism, while in this case the SR paradigm is applied, which avoids the introduction of per-path state into the network.

In general this approach applies to both IS-IS and OSPF, while the specific protocol extensions and encodings are different. In the current version of this document, the required IS-IS extensions are described. The required OSPF extensions will be described in a future version.

4. SR Virtual Topology with Resource Guarantee

As described in [I-D.ietf-isis-segment-routing-extensions], the IS-IS TLV-222 (MT-ISN) and TLV-223 (MT IS Neighbor Attribute) have been enhanced to carry the Adj-SID sub-TLV, and TLV-235 (Multitopology IPv4 Reachability) and TLV-237 (Multitopology IPv6 IP Reachability) have been enhanced to carry the Prefix-SID sub-TLV. With these enhancements, dedicated Segment Identifiers (SIDs) can be assigned for each SR virtual network topology.

This section specifies the necessary extensions to enable the deployment of resource guaranteed SR virtual topologies. Each virtual topology can be allocated with a particular partition of network resources from the underlay network, the SIDs associated with each SR virtual topology are used to identify the set of resources allocated from the network elements.

4.1. Topology specific Link Resource Allocation and Identification

A network link can participate in one or multiple SR virtual topologies, each virtual topology is assigned with a dedicated adj-SID. In order to describe the amount of link resource allocated to a

particular SR virtual topology, a new IS-IS sub-TLV called "SR Bandwidth" sub-TLV is defined:

The SR-Bandwidth sub-TLV is an optional sub-TLV carrying the aggregated bandwidth allocated to a particular SR adj-SID, which is associated witha particular virtual topology. In the data plane, the allocated bandwidth and the associated functional components are identified by the adj-SID of the virtual topology. This sub-TLV may be advertised as a sub-TLV of the following TLVs:

TLV-22 (Extended IS reachability) [RFC5305]

TLV-23 (IS Neighbor Attribute) [RFC5311]

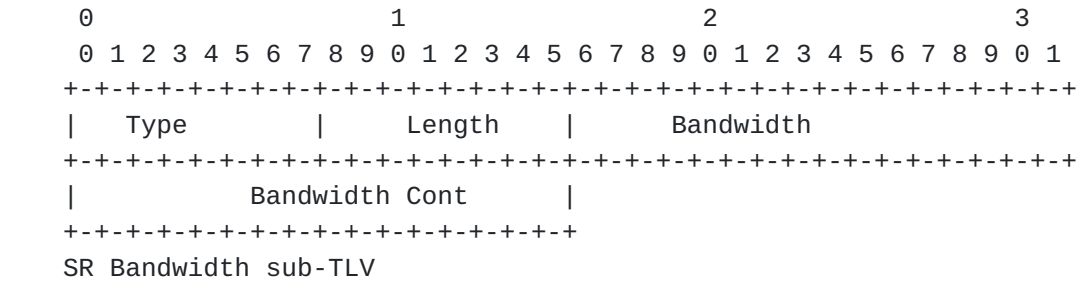
TLV-141 (inter-AS reachability information) [RFC5316]

TLV-222 (Multitopology IS)[RFC5120]

TLV-223 (Multitopology IS Neighbor Attribute) [RFC5311]

The SR bandwidth sub-TLV can appear at most once for a particular topology. Multiple SR Bandwidth sub-TLVs MAY be associated with a single IS neighbor.

The following format is defined for the SR Bandwidth sub-TLV:



SR Bandwidth sub-TLV

where:

Type: TBD, to be assigned by IANA.

Length: variable.

The SR bandwidth is encoded in 32 bits in IEEE floating point format. The units are bytes (not bits!) per second.

[I-D.ietf-teas-sr-rsvp-coexistence-rec] describes several options for traffic engineering in networks where RSVP-TE and SR LSPs coexist. Note that section 3.1 of [I-D.ietf-teas-sr-rsvp-coexistence-rec] proposes to partition the network bandwidth between RSVP-TE and SR.

The can be considered as a special case of creating one default SR virtual topology with dedicated bandwidth allocated, so that the network resources and operation of SR are isolated from the RSVP-TE based LSPs.

4.2. Topology specific Node Resource Allocation and Identification

A network node can participate in one or multiple SR virtual topologies, each virtual topology is assigned with a dedicated node-SID. In SR loose path forwarding, the topology specific node-SIDs can be used by transit network nodes to identify the virtual topology the packet belongs to, so as to steer the packet through the set of link resources allocated for the identified virtual topology. A prefix-SID sub-TLV describing the dedicated node-SID for each virtual topology is needed, this is supported in [I-D.ietf-isis-segment-routing-extensions].

In addition, similar to the allocation of link resource to virtual topologies, it is possible to allocate a subset of nodal resources for a particular virtual topology to ensure end-to-end service delivery. The nodal resources can be identified by the topology specific node-SIDs, so that in data plane the node SIDs can be used to steer a packet through the set of nodal resources allocated to this topology. Optional sub-TLVs describing the allocated resources at the node level for a particular virtual topology may be defined in future. The specification of nodal resources is for further study.

5. Multiple Services in SR Virtual Topology

Within one SR virtual topology, one or more types of service can be deployed using the resources allocated to this virtual topology. Each service type can have specific resource constraints and characteristics. The concept is similar to the DS-TE model [RFC4124] of RSVP-TE based mechanism, while in this case the SR paradigm is applied, which avoids the introduction of per-flow state into the network.

Some mechanism is needed to identify different service types and specify the different service characteristics within one virtual topology. The detailed protocol extensions will be provided in a future version.

5.1. Common Service Types

A service type is fundamentally the sum of the properties of a group of services. The authors considered specifically creating a number of specific service types within the protocol but concluded that this

was meaningless. The following sections show how a number of well known service types can be constructed.

5.1.1. Best Effort

Best effort service can be the only service type in a particular virtual topology. In this case, all of the resources allocated to this virtual topology instance are available to the best effort services.

Where there are multiple service types being carried in a virtual topology, best effort service will be transmitted over the links and nodes when there is an opportunity. The maximum resources which can be used by best effort service may be constrained to a subset of the topology resource. The Traffic Class (TC) of the best effort service SHOULD be set to lower than any other service types.

In the data plane, the SID and the Traffic Class value in the packet can be used to identify the service type and steer the best effor packets into the correct forwarding resources, such as queues.

Best effort services may or may not be protected at the discretion of the network operator.

5.1.2. Assured Bandwidth

An Assured Bandwidth service is one in which the bandwidth is assured but the latency is not. Thus, some bandwidth can be allocated to the assured bandwidth service, and traffic up to that bandwidth will be transmitted over the service, but the traffic may be delayed by other traffic.

It is likely that the assured bandwidth service will be carried in a virtual topology together with other service types, such as the best effort service. The maximum resources which can be used by assured bandwidth service SHOULD be constrained to a subset of the topology resource.

There will frequently be more than one assured bandwidth service running on a topology, and the Traffic Class (TC) could be used to determine how the various services compete for access to the link. Whilst the bandwidth is assured over the long term, over the short term it is not and such services will interact with similar and lower service classes in such a way that packet delay and jitter is not assured.

In the data plane, the SID and the Traffic Class value in the packet can be used to identify the service type and priority and steer the

assured bandwidth service packets into the correct forwarding resources, such as queues.

5.1.3. Deterministic

A Deterministic service is a service that may have controlled delay/jitter characteristics and/or an enhanced packet delivery assurance. Delay/Jitter may be addressable through the provision of sufficient bandwidth or it may require some form of packet scheduling. Enhanced delivery assurance may require the use of packet replication and elimination mechanism. The design of a deterministic network is discussed in [I-D.ietf-detnet-architecture]. Note that delay protection and delivery protection are orthogonal characteristics and a service may provide just one of the characteristics or it may provide both.

The details of a deterministic service will be provided in a future version. Such a service may be specified using the TLVs defined in [I-D.geng-detnet-info-distribution]

6. Topology and Algorithm

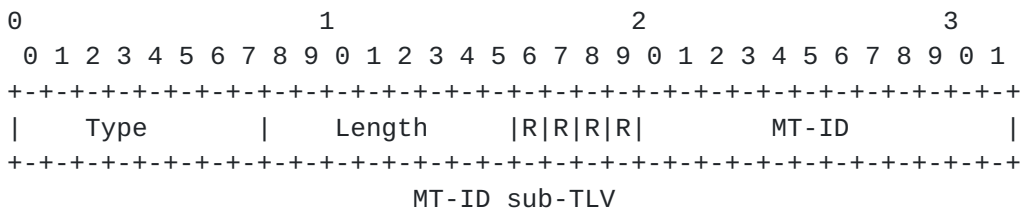
In the proposed mechanism, SR is used with IGP multi-topology to create one or more SR virtual topologies, each associated with a set of network resources allocated for the virtual topology. The service paths used between nodes in one virtual topology are not constrained to be shorted path by IGP metric, and can be any non-looping path that best suits the needs of the service. These paths may be imposed by the network controller, or calculated using a distributed method. For example, different SR algorithms as defined in [I-D.ietf-isis-segment-routing-extensions] can be used within one virtual topology. The flex-algo mechanism defined in [I-D.ietf-lsr-flex-algo] may also be used in one virtual topology to meet different service requirements.

7. SRv6 Considerations

The mechanisms to create virtual network topologies with allocated resources using an SRv6 data-plane are similar to SR with an MPLS data plane, although there are some differences to be considered. This section specifies the necessary protocol extensions to enable SRv6 with multi-topology and the mechanisms defined in this document. Detailed method of operating enhanced VPN over an SRv6 data-plane will be described in a future version.

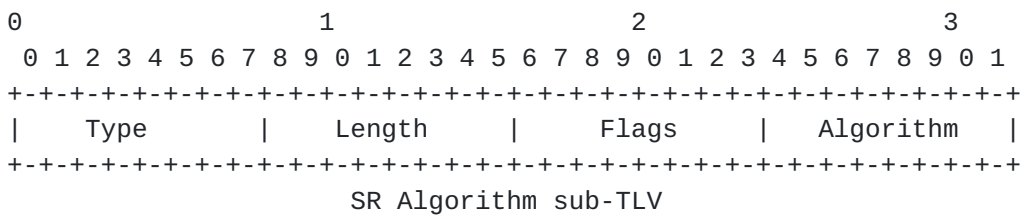
In [I-D.bashandy-isis-srv6-extensions], the SRv6 node SID TLV is defined as a top-level TLV, which cannot be carried under the MT IPv6 IP Reach TLV (type 237). In order to specify the association between

SRv6 node SIDs and differernt virtual topologies, a new sub-TLV called "MT-ID sub-TLV" under the SRv6 node SID TLV is introduced. The format of the sub-TLV is as follows:



MT-ID sub-TLV

In addition, in order to advertise multiple algorithms used for calculating reachability to nodes within a particular topology, a new sub-TLV called "SR algorithm sub-TLV" under the SRv6 node SID TLV is introduced. The format of the sub-TLV is as follows:



SR Algorithm sub-TLV

In order to advertise the SRv6 adj-SIDs associated with different topologies the network node participates in, the SRv6 Adjacency SID sub-TLV as defined in [I-D.bashandy-isis-srv6-extensions] MUST be carried in the MT Intermediate Systems TLV (type 222). The SR bandwidth sub-TLV as defined in this document SHOULD also be carried in the MT Intermediate Systems TLV.

8. Fast Repair

In some instances it is desirable to provide some form of fast repair for a failed link or node. The methods available fall into two categories, end-to-end, for example 1+1, and IP fast reroute. Which ever of these is used it is desirable that the repair path provides the same level of service to the tenant as the tenant's normal service. This would mean that the repair path needs to be constrained to the tenant's topology, or to some repair topology reserved exclusively for that tenant for the duration of the repair. The normal way that IPFRR operates is that the point of local repair (PLR) calculates the repair path based on the information flooded by the routing protocol. How the PLR can maintain the level of service through the repair is for further study.

9. LAN interface

The use of multi-point to multi-point (MP2MP) interfaces is currently out of scope for this design.

A LAN interface MUST be used in point to point mode.

Note support for MP2MP may be needed in the future, and this is for further study.

10. Security Considerations

This document introduces no additional security vulnerabilities to IS-IS and OSPF.

The mechanism proposed in this document is subject to the same vulnerabilities as any other protocol that relies on IGPs.

11. IANA Considerations

This document requests IANA to allocate a sub-TLV type as defined in Section 4 from "Sub-TLVs for TLVs 22, 23, 25, 141, 222 and 223" registry.

Value	Description	Reference
TBA1	SR bandwidth sub-TLV	This document

Per TLV information where SR bandwidth sub-TLV can be part of:

TLV	22	23	25	141	222	223
	y	y	n	y	y	y

This document requests IANA to allocate 2 sub-TLVs type as defined in Section 7 from the "Sub-TLVs for TLVs 27, 135, 235, 236 and 237" registry.

Value	Description	Reference
TBA2	MT-ID sub-TLV	This document
TBA3	SR Algorithm sub-TLV	This document

Per TLV information where the sub-TLVs can be part of:

TLV	27	135	235	236	237
TBA2	y	n	n	n	n
TBA3	y	n	n	n	n

12. Acknowledgments

The authors would like to thank Mach Chen and Robin Li for the review and discussion of this document.

13. References

13.1. Normative References

[I-D.dong-spring-sr-for-enhanced-vpn] Dong, J., Bryant, S., Li, Z., and T. Miyasaka, "Segment Routing for Enhanced VPN Service", draft-dong-spring-sr-for-enhanced-vpn-00 (work in progress), March 2018.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <https://www.rfc-editor.org/info/rfc2460>.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

[RFC4915] Psenak, P., Mirtorabi, S., Roy, A., Nguyen, L., and P. Pillay-Esnault, "Multi-Topology (MT) Routing in OSPF", RFC 4915, DOI 10.17487/RFC4915, June 2007, <https://www.rfc-editor.org/info/rfc4915>.

[RFC5120] Przygienda, T., Shen, N., and N. Sheth, "M-ISIS: Multi Topology (MT) Routing in Intermediate System to Intermediate Systems (IS-ISs)", RFC 5120, DOI 10.17487/RFC5120, February 2008, <https://www.rfc-editor.org/info/rfc5120>.

[RFC5311] McPherson, D., Ed., Ginsberg, L., Previdi, S., and M. Shand, "Simplified Extension of Link State PDU (LSP) Space for IS-IS", RFC 5311, DOI 10.17487/RFC5311, February 2009, <https://www.rfc-editor.org/info/rfc5311>.

[RFC7810] Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 7810, DOI 10.17487/RFC7810, May 2016, <https://www.rfc-editor.org/info/rfc7810>.

13.2. Informative References

[I-D.bashandy-isis-srv6-extensions]
Ginsberg, L., Bashandy, A., Filsfils, C., and B. Decraene, "IS-IS Extensions to Support Routing over IPv6 Dataplane", draft-bashandy-isis-srv6-extensions-01 (work in progress), September 2017.

[I-D.bryant-rtgwg-enhanced-vpn]
Bryant, S. and J. Dong, "Enhanced Virtual Private Networks (VPN+)", draft-bryant-rtgwg-enhanced-vpn-01 (work in progress), October 2017.

[I-D.geng-detnet-info-distribution]
Geng, X. and M. Chen, "IGP-TE Extensions for DetNet Information Distribution", draft-geng-detnet-info-distribution-01 (work in progress), September 2017.

[I-D.ietf-detnet-architecture]
Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", draft-ietf-detnet-architecture-04 (work in progress), October 2017.

[I-D.ietf-isis-segment-routing-extensions]
Previdi, S., Ginsberg, L., Filsfils, C., Bashandy, A., Gredler, H., Litkowski, S., Decraene, B., and J. Tantsura, "IS-IS Extensions for Segment Routing", draft-ietf-isis-segment-routing-extensions-15 (work in progress), December 2017.

[I-D.ietf-lsr-flex-algo]
Psenak, P., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", draft-ietf-lsr-flex-algo-00 (work in progress), May 2018.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-15 (work in progress), January 2018.

[I-D.ietf-teas-sr-rsvp-coexistence-rec]
Sitaraman, H., Beeram, V., Minei, I., and S. Sivabalan, "Recommendations for RSVP-TE and Segment Routing LSP co-existence", draft-ietf-teas-sr-rsvp-coexistence-rec-04 (work in progress), May 2018.

[RFC4124] Le Faucheur, F., Ed., "Protocol Extensions for Support of Diffserv-aware MPLS Traffic Engineering", RFC 4124, DOI 10.17487/RFC4124, June 2005, <https://www.rfc-editor.org/info/rfc4124>.

Authors' Addresses

Jie Dong
Huawei Technologies

Email: jie.dong@huawei.com

Stewart Bryant
Huawei Technologies

Email: stewart.bryant@gmail.com